



Aslib Journal of Information Management

Geographic variation in social media metrics: an analysis of Latin American journal articles

Juan Pablo Alperin

Article information:

To cite this document:

Juan Pablo Alperin , (2015), "Geographic variation in social media metrics: an analysis of Latin American journal articles", Aslib Journal of Information Management, Vol. 67 Iss 3 pp. 289 - 304

Permanent link to this document:

<http://dx.doi.org/10.1108/AJIM-12-2014-0176>

Downloaded on: 01 November 2016, At: 23:04 (PT)

References: this document contains references to 39 other documents.

To copy this document: permissions@emeraldinsight.com

The fulltext of this document has been downloaded 333 times since 2015*

Users who downloaded this article also downloaded:

(2015), "The thematic orientation of publications mentioned on social media: Large-scale disciplinary comparison of social media metrics with citations", Aslib Journal of Information Management, Vol. 67 Iss 3 pp. 260-288 <http://dx.doi.org/10.1108/AJIM-12-2014-0173>

(2015), "Usefulness of altmetrics for measuring the broader impact of research: A case study using data from PLOS and F1000Prime", Aslib Journal of Information Management, Vol. 67 Iss 3 pp. 305-319 <http://dx.doi.org/10.1108/AJIM-09-2014-0115>

Access to this document was granted through an Emerald subscription provided by emerald-srm:563821 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Geographic variation in social media metrics: an analysis of Latin American journal articles

Social media
metrics

289

Juan Pablo Alperin

*Publishing Program & Public Knowledge Project,
Simon Fraser University, Vancouver, Canada*

Received 19 December 2014

Revised 18 March 2015

Accepted 24 March 2015

Abstract

Purpose – The purpose of this study is to contribute to the understanding of how the potential of altmetrics varies around the world by measuring the percentage of articles with non-zero metrics (coverage) for articles published from a developing region (Latin America).

Design/methodology/approach – This study uses article metadata from a prominent Latin American journal portal, SciELO, and combines it with altmetrics data from Altmetric.com and with data collected by author-written scripts. The study is primarily descriptive, focusing on coverage levels disaggregated by year, country, subject area, and language.

Findings – Coverage levels for most of the social media sources studied was zero or negligible. Only three metrics had coverage levels above 2 per cent – Mendeley, Twitter, and Facebook. Of these, Twitter showed the most significant differences with previous studies. Mendeley coverage levels reach those found by previous studies, but it takes up to two years longer for articles to be saved in the reference manager. For the most recent year, coverage was less than half than what was found in previous studies. The coverage levels of Facebook appear similar (around 3 per cent) to that of previous studies.

Research limitations/implications – The Altmetric.com data used for some of the analyses were collected for a six month period. For other analyses, Altmetric.com data were only available for a single country (Brazil).

Originality/value – The results of this study have implications for the altmetrics research community and for any stakeholders interested in using altmetrics for evaluation. It suggests the need of careful sample selection when wishing to make generalizable claims about altmetrics.

Keywords Social media, Latin America, Metrics, Altmetrics, Scholarly publishing, SciELO

Paper type Research paper

Introduction

Altmetrics hold the potential to change how research is discovered, disseminated, evaluated, rewarded, and even read. They could become a transformative force in academia, but only if they are shown to actually provide “indications of impacts on diverse audiences including scholars but also practitioners, clinicians, educators and the general public” (Piwowar, 2013, p. 9) and to present the opportunity to measure and track additional aspects of impact – aspects ignored by traditional bibliometric approaches. However, to date, altmetrics remain largely understudied and often misunderstood.

The perceived potential of altmetrics, coupled with this lack of understanding of the behaviours that give rise to the metrics, has turned altmetrics into a growing field of

© Juan Pablo Alperin 2015

This research was supported by the project “Quality in the Open Scholarly Communication of Latin America”: IDRC Grant reference number: 106660-001. The author also wishes to thank Euan Adie (Altmetric.com) for granting access to Altmetric.com data, and Fabio Batalha (SciELO Brazil) for tolerating countless e-mails and phone calls asking for the SciELO download data to be updated.



scholarly inquiry. In the process, several definitions have emerged that propose, in varying levels of detail, that altmetrics are essentially about examining the social web as an alternative or enhancement to traditional bibliometrics (Galligan and Dyas-Correia, 2013). Unfortunately, this reliance on the social web means that any relationships or conclusions drawn from altmetrics are fraught with problems of generalizability. Unlike citations, which are drawn from curated databases with known properties, altmetrics' reliance on the web means that the metrics must be disentangled from the underlying usage patterns of the web's chaotic social services. This is particularly problematic given the "staggering amount of unevenness in the production and dissemination of information [on the Web]" (Graham *et al.*, 2011, p. 26)[1].

This unevenness must be taken into account in the search for a generalized understanding of what altmetrics mean and what they can offer. There are on-going efforts, such as the NISO Altmetrics Project (NISO, 2014), that aim to standardize the metrics, but these efforts must necessarily be informed by studies that look at altmetrics in journals and articles published and read in a broad range of contexts, including research that is published and read in regions of the world beyond the global North[2], as well as in other languages beyond English. Only by studying altmetrics in a diverse set of contexts (e.g. by geography, language, content) will it be possible to understand the limits and potential of what altmetrics can measure.

Given the relative recency of altmetrics, further work is still needed in all contexts. This study seeks to contribute to the literature of the altmetrics community by being the first to explore the coverage levels between countries, subjects, and languages of articles published from Latin America. In doing so, it focuses special attention to the region's largest country: Brazil. By exploring a developing region broadly and an emerging country in detail, this research hopes to open a window on the universality and limits of some of the existing claims from the altmetrics literature, and through that window, provide a clearer understanding of the differences in the potential of altmetrics around the world.

Background

Altmetrics work by looking for references to scholarly works on the web, including "traditional" social media (i.e. Twitter, Facebook, Google+), blogs (i.e. researchblogging.com, ScienceSeeker, Wordpress.com), academic bookmarking services and reference managers (i.e. CiteULike, Mendeley, Connotea), media outlets (i.e. *New York Times*, *The Economist*, *Wired*), and multimedia (i.e. YouTube, podcasts), post-publication peer review sites (i.e. F1000 Prime), and a handful of others. However, there is no official list of what constitutes an alternative metric. Virtually any metric that can be collected over the web, and that is not a citation, can be considered an alternative metric. One of the most complicated aspects of altmetrics is that each of these sources has a different degree of adoption and use around the world and between different online communities. An understanding of this difference in use must necessarily influence the metrics that can be derived from each source.

Even in the face of disparate use across contexts, altmetrics warrant attention, not least because of the growing importance of the role of the web in the communication of science. For example, it is thought that Twitter use among scholars has been growing (Priem *et al.*, 2012), and several surveys confirm that most scholars are making at least some use of "Web 2.0" tools (Procter *et al.*, 2010, Tenopir *et al.*, 2013)[3]. More generally, the number of articles that are shared or discussed each day is increasing at an

estimated rate of between 5 and 10 per cent per month (Adie and Roe, 2013). So while the presence of articles in social media is still relatively low – at an estimated 15-24 per cent (Costas *et al.*, in press) – the potential of these metrics increases daily, and with them the need to understand altmetrics across a wide range of contexts.

Even as there is more scrutiny on mentions of articles on the social web and as coverage levels are better understood, it is still unclear to what degree altmetrics can be used for capturing different forms of impact. There is a definite sense that the numbers in some way indicate attention, influence, or even impact, but it is still unknown how. As a consequence, most of the existing altmetric studies so far have focused on the relationship between altmetrics and citation metrics. So far, the strongest correlation between any altmetrics source and citations were found for F1000 Prime recommendations (Li and Thelwall, 2012). However, several studies have shown a moderate level of correlation between saves in the bibliographic manager Mendeley and citations (Li *et al.*, 2012; Priem *et al.*, 2012; Bar-Ilan, 2012; Bar-Ilan *et al.*, 2012). Mentions in blogs has also shown to have a positive correlation with citations (Fausto *et al.*, 2012; Costas *et al.*, in press), although this appears to be heavily influenced by the makeup of bloggers themselves and the journals they tend to blog about, both of which introduce a strong bias for high-impact life science journals (Shema *et al.*, 2012). Twitter, the most prevalent of the social media altmetrics sources (Thelwall *et al.*, 2013), was found to have a low (but positive) correlation to citations (Haustein *et al.*, 2014c; Eysenbach, 2011). Eysenbach (2011) additionally found that, for a small sample of articles from a single journal, highly tweeted papers could also be used as early predictors of citations, something supported by the findings of Shuai *et al.* (2012) and Thelwall *et al.* (2013). In short, there appears to be a low to moderate correlation between some altmetric sources and citations, at least in the specific journals and disciplines studied (primarily well-known English-language natural and life science journals)[4].

While the presence of scholarly articles in online channels of communication is growing, most research on altmetrics conclude with a call for further study of the reliability, validity, and context of the available metrics (Haustein *et al.*, 2014c; Wouters and Costas, 2012; Liu and Adie, 2013). The existing studies have primarily used well-established journals such as *Science* and *Nature* (Li *et al.*, 2012), and *JASIST* (Bar-Ilan, 2014) or large archives such as arXiv.org and PubMed (Shuai *et al.*, 2013; Haustein *et al.*, 2014a, c), or in the broadest of cases, samples drawn from bibliographic databases like the web of science (WoS) (Zahedi *et al.*, 2014a)[5]. This limited array of sources introduces a selection bias that calls into question the generalizability of reported results. The appeal of these sources is obvious – they are prominent and their data are easily available for analysis – but there is a need to turn to a different geographic, economic, and socio-cultural context in order to understand the universality of claims about altmetrics. Latin America's leading research publishing initiatives provide such an alternative context. This study therefore presents an exploration of the penetration of altmetrics by looking at a large Latin American scholarly journal portal.

There are many scholarly publishing initiatives in Latin America. They range from small and institutional, to large and regional; from broad multidisciplinary to subject specific; from those focused on scholarly journals, to those focused on institutional archiving; and they can be found in virtually every country between Mexico and Argentina. The number of initiatives continues to grow, but the three major initiatives – Latindex, RedALyC, and SciELO – are arguably the most influential and certainly provide the best available data on the scholarly publishing activities of the region (Alperin, 2014a, b).

Data and methods

This study focuses on SciELO for access to both articles and journal metadata. SciELO publishes and indexes full-text articles of over 1,200 journals and over 500,000 full-text articles across 12 national collections from Iberomaerica and the Caribbean[6]. While SciELO is technically a publisher (i.e. it publishes content online which in some cases has not been made available elsewhere), it is not a publisher in the traditional sense (i.e. it is not involved in the operation of the journals, nor does it provide financing, editing, or other oversight of the editorial or production process). It has been described as a both a “meta-publisher” (Packer and Meneghini, 2007) and as a hybrid between a repository and a publisher (Guédon, 2008).

SciELO is widely recognized in the region (Gómez *et al.*, 2013) and has gained acceptance as a marker of quality by several national science councils (Alperin *et al.*, 2011) because it hosts some of the best-regarded journals published from within Latin America. This study focuses on journals from the nine national collections from Latin America and the Caribbean, namely Argentina, Brazil, Chile, Colombia, Costa Rica, Cuba, Mexico, Peru, and Venezuela[7]. These collections differ greatly in size, with the flagship collection (Brazil), publishing over four times as many articles as the next largest collection (Chile) (Table I). Not only does SciELO provide access to full-text articles through their web portal, it also provides curated metadata (bibliographic information), download data, and citation data for the articles they publish.

Article-level data from SciELO were collected through available application programming interfaces (APIs) for the years starting in 2000[8]. The data were last updated in May 2014, and includes articles that were available by that date, for a total of 389,795 articles from the Latin American portals, with Brazil making up 58 per cent of those articles. The corpus covers a wide range of subject fields in varying degrees (Table II). It is difficult to compare coverage levels across subjects because, like in the WoS, journals can be categorized in more than one field simultaneously, but unlike WoS, SciELO does not provide a single top-level categorization. SciELO’s origins in the Health Sciences are evident from the fact that over 40 per cent of the articles are classified in that category. The next most prevalent categories are the Humanities with 20 per cent, followed by Agricultural Sciences, Applied Social Sciences, and Biological Sciences each around 12-15 per cent[9].

While detailed metadata, download statistics, and citation data for the articles were downloaded from SciELO, most of the altmetric data were provided by the largest of altmetric providers, Altmetric.com, including the number of mentions in social media

Country	Number of journals	Number of articles 2013	Number of articles since 2000
Argentina	82	2,212	21,507
Brazil	282	21,880	224,292
Chile	94	4,285	38,988
Colombia	161	4,749	35,736
Costa Rica	13	486	4,592
Cuba	46	2,356	19,535
Mexico	103	3,342	24,205
Peru	16	663	5,388
Venezuela	20	456	15,552
Total	817	40,429	389,795

Table I.
Number of journals
and articles in
SciELO collection
by country

(i.e. Twitter, Facebook, Google+), videos (i.e. youtube.com), and mainstream media (i.e. NYTimes, Slate)[10]. The data were provided in June 2014 and comprise all mentions of any URLs within the SciELO domains (i.e. containing “scielo” in the domain name) in the collected altmetrics sources. For most SciELO collections, the Altmetric.com data goes back to November 2013, but in the case of SciELO Brazil (scielo.br) an extended set goes back to January 2013[11].

The article data from SciELO were linked with the Altmetric.com data by matching on the internal SciELO article ID. The article ID were immediately available in the SciELO data set, and was extracted from the Altmetric.com data set by looking at the URLs mentioned in every record. The SciELO ID can be seen in the article’s URL in the form of a URL parameter, making mentions of SciELO articles (and their corresponding ID) unambiguously identifiable. This process lead to the identification of 8,427 total mentions (in any source) across 3,647 SciELO articles (in any collection). For the extended Altmetric.com data set corresponding to only scielo.br articles from 2013, the data were linked to the scielo.br data using a DOI available in both data sets, yielding a total of 3,588 mentions across 1,686 articles.

In addition, a series of Python scripts (written by the author) were used to query the number of times articles are mentioned in the English, Spanish, and Portuguese editions of Wikipedia and saved in the Mendeley reference manager. Data from Mendeley were fetched by searching the Mendeley API with the article title, and verifying the search results using the year of publication and the first author last name [12]. Similarly, Wikipedia counts were fetched by searching the Wikipedia API for mentions of partial URLs, which were specific enough to uniquely identify mentions of SciELO articles, but general enough to catch all URL variants of an article[13]. Both scripts were run in September 2014. The same method described for the Altmetric.com data were used to extract and link the SciELO ID from the URLs identified on Wikipedia.

The resulting data set contains every SciELO article published in Latin America between January 2000 and May 2014 linked using the SciELO ID to every mention in the Altmetric.com data set that occurred since November 2013. The data set also includes any mention of those articles in Mendeley and in English, Spanish, and Portuguese versions of Wikipedia as of September 2014.

Results

With the data in place, it is possible to measure coverage across the entire corpus, between country, disciplines, and language in two ways: by looking at the total number

Subject	Number of journals 2013	Number of articles 2013
Agricultural sciences	89	5,681
Health sciences	242	16,607
Applied social sciences	200	5,093
Human Sciences	263	8,098
Exact and Earth Sciences	67	2,502
Engineering	77	2,908
Biological Sciences	101	5,039

Notes: Journals may be classified in more than one category. A total of 4,542 articles are classified in at least two categories, and 496 of those are classified in at least four categories

Table II.
Number of journals
and articles in all
SciELO collections
by subject

of social media mentions, and by looking at the proportion of articles that receive mentions. The raw counts give a sense of the scale of activity surrounding the articles (i.e. how many people are engaging with these articles on social media) and the percentage coverage indicates how concentrated or spread out the activity is across all articles (i.e. how many articles solicit engagement).

The first thing to note is that across the SciELO collections, the source with the most mentions (by several orders of magnitude) is Mendeley, with Twitter being a distant second, and Facebook third (Table III). Other sources have negligible coverage or none at all[14]. The total of all mentions in the data set is 1,189,583, of which Mendeley makes up almost 80 per cent and Twitter just over 8per cent[15]. These mentions are associated with 173,733 articles (Table IV).

As can be seen in Tables III and IV, the Brazilian journals on SciELO have far larger number of mentions than any other country's journals. It is worth repeating that for all sources except Mendeley and Wikipedia, numbers are based only on mentions that occurred between 1 November 2013 and 31 May 2014, a period that typically sees lower overall use because it spans the Christmas and the summer holidays in the Southern hemisphere, as well as Carnival in Brazil. When Mendeley is excluded from the analysis, the number of mentions drops drastically to only 9,372 mentions spanning 4,399 articles (just over 1 per cent of all articles).

The immediacy and obsolescence of most altmetrics has still been relatively understudied, but there are clear signs that, like with citations-based metrics, time since

Table III.
Total number
of mentions

Country	Blogs	Facebook	Google +	Mass Media	Twitter	Videos	Wikipedia (en, es, pt)	Mendeley	All mentions	All mentions (without Mendeley)
Argentina	-	39	-	-	411	3	64	27,755	27,755	517
Brazil	6	735	93	7	4,033	57	427	936,163	936,163	5,358
Chile	-	34	6	-	1,388	3	303	95,102	95,102	1,734
Colombia	-	7	2	-	273	1	26	38,033	38,033	309
Costa Rica	1	-	-	1	65	-	12	16,308	16,308	79
Cuba	-	12	4	-	327	2	17	11,361	11,361	362
Mexico	-	22	2	-	332	-	42	33,407	33,407	398
Peru	-	3	1	-	117	-	17	11,319	11,319	138
Venezuela	-	4	2	-	434	-	37	10,763	10,763	477
Total	7	856	110	8	7,380	66	945	1,180,211	1,180,211	9,372

Table IV.
Total number of
document mentions

Country	Number of articles	Blogs	Facebook	Google +	Mass media	Twitter	Videos	Wikipedia (en, es, pt)	Mendeley	Any Altmetric
Argentina	21,507	-	32	-	-	216	3	57	3,204	3,465
Brazil	224,292	6	551	32	6	1,402	36	369	134,849	135,848
Chile	38,988	-	31	5	-	568	3	236	16,700	17,130
Colombia	35,736	-	4	2	-	136	1	21	6,131	6,254
Costa Rica	4,592	1	-	-	1	44	-	10	1,555	1,598
Cuba	19,535	-	7	4	-	208	2	17	1,954	2,138
Mexico	24,205	-	18	2	-	207	-	37	4,605	4,797
Peru	5,388	-	3	1	-	51	-	11	742	791
Venezuela	15,552	-	4	2	-	187	-	33	1,513	1,712
Total	389,795	7	650	48	7	3,019	45	791	171,253	173,733

publication is an important measure. In several multi-year studies, most altmetrics show an increasing presence in more recent publications (Costas *et al.*, in press; Zahedi *et al.*, 2014a; Haustein *et al.*, 2014b), although Zahedi *et al.* (2014a) show that when looked at as a percentage of all articles published, some sources (such as Mendeley) tend to peak a few years prior to the current year (likely indicative of a multi-year obsolescence period), whereas other sources (such as Twitter) have higher coverage as a percentage of all articles for the most recent years (likely indicative of increasing adoption coupled with a shorter obsolescence period).

SciELO articles do not deviate from this pattern. Articles display the expected pattern of higher mentions for more recent publications, with Mendeley having higher coverage as a percentage of all articles a few years prior (peak in 2007/2008), although 2012 already shows twice the coverage levels of 2013. The apparent short obsolescence period for most of the metrics studied suggests that for a more meaningful analysis across multiple metrics it would be more fruitful to focus on metrics that were collected closer to the publication date. The extended set of Altmetric.com data that includes SciELO Brazil (scielo.br) data dating back to January 2013 suits this purpose (all other characteristics of the data remain the same). However, the significantly higher coverage of Mendeley for previous years (up to three times that of 2013) suggests that it may be worthwhile to study Mendeley on its own (Table V).

The number of scielo.br articles from 2013 with social media metrics is significantly higher than for any other country in that year, although the numbers are quite low even here. Across sources, 5,319 (25.09 per cent) of articles had at least one metric, but if Mendeley is excluded, only 1,701 (8.02 per cent) have at least one metric. For all of 2013, several of the categories did not have any mentions (LinkedIn, Reddit, Pinterest, Q&A, peer review). Other categories also had < 0.01 per cent coverage (< 25 articles) (news, blogs, Google+, F1000, video, and Wikipedia). The only three sources with any significant number of mentions were Mendeley, Twitter, and Facebook.

Table VI summarizes the levels of coverage for these three main metrics at the article level. There were 3,986 (18.80 per cent) articles with at least one Mendeley reader,

Year	Facebook	Twitter	Mendeley	Wikipedia (en, es, pt)	Any Altmetric (except Mendeley)
2000	22 (0.3%)	31 (0.5%)	4,258 (62.8%)	17 (0.3%)	111 (1.6%)
2001	12 (0.2%)	24 (0.3%)	5,007 (63.8%)	20 (0.3%)	108 (1.4%)
2002	15 (0.2%)	27 (0.3%)	6,205 (66%)	24 (0.3%)	139 (1.5%)
2003	18 (0.2%)	23 (0.2%)	7,400 (67.2%)	18 (0.2%)	137 (1.2%)
2004	23 (0.2%)	39 (0.3%)	8,488 (68.5%)	39 (0.3%)	193 (1.6%)
2005	26 (0.2%)	23 (0.2%)	9,561 (69.3%)	35 (0.3%)	212 (1.5%)
2006	28 (0.2%)	72 (0.5%)	10,648 (69.8%)	28 (0.2%)	279 (1.8%)
2007	38 (0.2%)	62 (0.4%)	12,109 (70.5%)	25 (0.1%)	247 (1.4%)
2008	32 (0.2%)	54 (0.3%)	12,868 (70.7%)	28 (0.2%)	297 (1.6%)
2009	30 (0.2%)	62 (0.3%)	13,392 (68.9%)	25 (0.1%)	305 (1.6%)
2010	47 (0.2%)	82 (0.4%)	13,536 (65.2%)	26 (0.1%)	355 (1.7%)
2011	53 (0.2%)	86 (0.4%)	13,264 (60.4%)	17 (0.1%)	366 (1.7%)
2012	60 (0.3%)	110 (0.5%)	12,258 (55.4%)	49 (0.2%)	451 (2%)
2013	135 (0.6%)	433 (2%)	4,017 (18.4%)	17 (0.1%)	872 (4%)

Table V.
Number of mentions
by source and
publication year

1,286 (6.03 per cent) with at least one Tweet, and 596 (2.81 per cent) articles with a Facebook public wall post. As discussed later, these levels of coverage are lower than what has been reported by other studies.

The levels of coverage vary by journal, with a handful of journals exhibiting nearly universal coverage. On average, 15.3 per cent of a journal's articles had at least one Mendeley reader (SD 20.0), 5.8 per cent of journal's articles had at least one tweet (SD 9.3), and 3.14 per cent of a journal's articles had at least one Facebook post (SD 4.5). The spread of coverage can be seen in Figure 1. In all three cases, there are numerous journals that extend well beyond the 75 per cent percentile, potentially indicative of a marked difference in user communities or potentially a different promotion and diffusion strategies employed by the journals.

The relationship between coverage and number of mentions in all three sources can be viewed by plotting the coverage levels against the mean number of mentions, which reveals a fairly strong correlation between the level of coverage a journal receives (y -axes) and the mean number of saves/mentions received by the articles in that journal (x -axes). This correlation is confirmed through an ordinary least-squares regression that results in R^2 of 0.69 for Mendeley, 0.77 for Twitter, and 0.82 for Facebook (Figure 2). This is perhaps unsurprising given that the mean number of mentions is so low in most cases, so each mention is likely the first for the article, simultaneously increasing the journal's coverage and mean number by one.

Unfortunately, the citation window in the scielo.br data set is too small to make comparisons to citations meaningful (there has not been sufficient time for citations to accumulate given that the Altmetric.com data are only available for articles published since 2013). In the absence of meaningful citation data, the top altmetrics were compared with the download counts of the full-text in the first 30 and 90 days after publication.

Table VI.
Summary of
Mendeley, Twitter,
and Facebook for
SciELO Brazil

Metric	Number of articles	Coverage (%)	Maximum	Mean	SD	Mean (when metric > 0)	SD (when metric > 0)
Mendeley	3,986	18.80	119	0.59	2.06	3.15	3.8
Twitter	1,286	6.07	50	0.12	0.83	2.05	2.72
Facebook	602	2.84	48	0.04	0.45	1.52	2.2

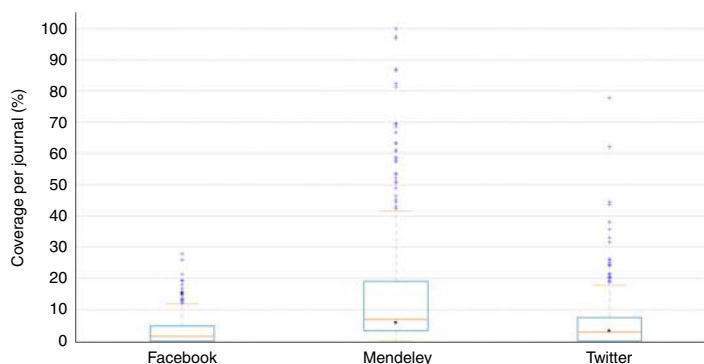


Figure 1.
Percentage of
articles with at least
one metric (coverage)
by journal and
platform (Facebook,
Mendeley, Twitter)

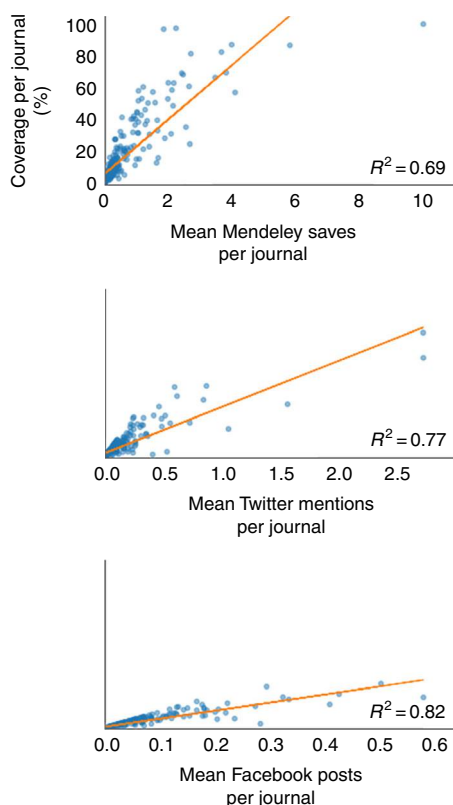


Figure 2.
Percentage of articles with at least one metric (coverage) by journal compared to mean number of mentions in Mendeley, Twitter, and Facebook

Mendeley, Tweets, Facebook, and number of full-text downloads in the first 30 and 90 days since publication were all correlated using Spearman's rank correlation (Table VII). Twitter was the most highly correlated of the altmetrics, with coefficients around 0.2 with both the download metrics and with Facebook. No correlation was found between Twitter and Mendeley. The Spearman coefficients for Facebook were a little lower (around 0.15 with downloads in first 30 and 90 days) and, as already noted, 0.2 with Twitter. Again, no correlation was found between posts Facebook and Mendeley readers. Mendeley showed the lowest correlations across the board, with correlations of around 0.1 with downloads. The correlations, where they exist, are

Metric	Twitter	Facebook	Mendeley	Downloads (first 30 days)	Downloads (first 90 days)
Twitter	1	0.21**	0.03**	0.22**	0.2**
Facebook	0.21**	1	0.01	0.15**	0.16**
Mendeley	0.03**	0.01	1	0.11**	0.13**
Downloads (first 30 days)	0.22**	0.15**	0.11**	1	0.84**
Downloads (first 90 days)	0.2**	0.16**	0.13**	0.84**	1

Note: **Correlation is significant at the 0.01 level (two-tailed)

Table VII.
Spearman rank correlation between Mendeley, Twitter, and Facebook

positive but low, similar to the effects found when comparing Tweets to citations (Haustein *et al.*, 2014c). As can be expected from these low correlations, there is little overlap in the sets of articles that receive Tweets, Facebook posts, and Mendeley saves (Figure 3).

Not only are the articles themselves different, but they are also comprised of different sets of languages, with coverage fluctuating depending on the language of the content. Content in English is about three times as likely to be saved in Mendeley than content in Spanish or Portuguese (29.0 per cent of English content has Mendeley readers, vs 11.2 and 9.6 per cent of Spanish and Portuguese, respectively). Content in Portuguese, however, is almost four times as likely to be shared on Facebook than content in Spanish, and almost twice as likely as content in English (3.8 per cent for Portuguese vs 1.0 per cent for Spanish and 1.7 per cent for English). It should be noted that, given that this data are drawn from the Brazilian journals in SciELO, the sample of Spanish articles is quite low ($n = 412$ articles), compared to the dominant Portuguese ($n = 11,909$) and popular English ($n = 8,850$).

Figure 4 shows the coverage of these three sources is on articles of different disciplines. The Humanities Sciences, for example, have among the lowest Mendeley coverage, but the highest Facebook coverage and second highest Twitter coverage. The converse is true of the Biological Sciences, which have the highest levels of Mendeley coverage but amongst the lowest Facebook coverage. Engineering, on the other hand, has very low coverage on social media, but moderate coverage in Mendeley. The Health Sciences seem to be the only field that is high across all three measures, with the highest or second highest coverage for each. These between-field fluctuations in coverage levels are indicative of differing uses of social media, potentially indicating different audiences are being reached by each discipline or different practices regarding social media and reference management.

Discussion and conclusions

The differences uncovered by this study point to the need of careful sample selection when making generalizable claims about altmetrics. In particular, it draws attention to the importance of considering and further studying altmetrics in developing and emerging country contexts.

The levels of coverage by the major altmetrics sources are generally lower than what has been previously reported. Most sources showed zero or almost zero coverage for all of the SciELO collection. Large differences were also found between the metrics that showed some coverage (Mendeley, Twitter, and Facebook).

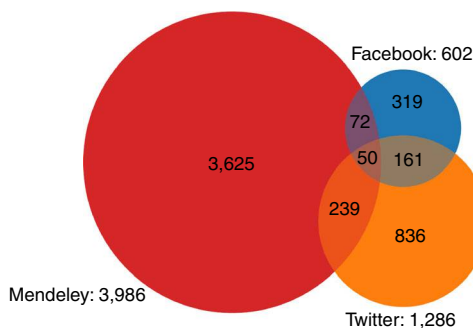


Figure 3.
Overlap of article
coverage between
Mendeley, Twitter
and Facebook

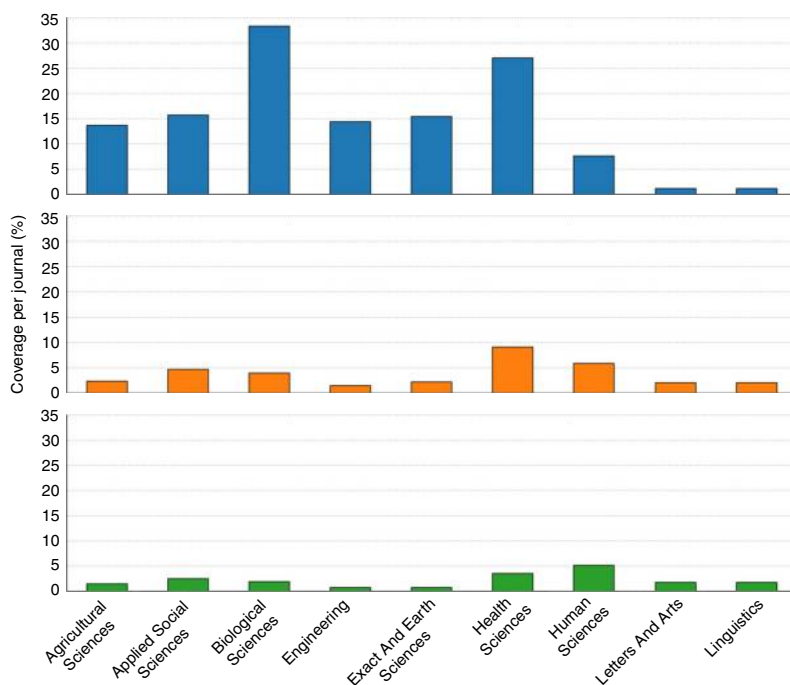


Figure 4.
Mendeley (top),
Twitter (middle), and
Facebook (bottom)
coverage by subject

Of these, the most dramatic differences were found for Twitter, for which Haustein *et al.* (2014c) found that over 20 per cent of articles from 2012 were tweeted at least once (twice that of 2011) using a set of articles in both PubMed and WoS. Other Twitter studies have found that 39 per cent of arXiv papers (2012) are tweeted at least once (Haustein *et al.*, 2014c) and over 11 per cent in an older study of 2010 PLOS One articles (Priem *et al.*, 2012). Again, the study showing the lowest coverage was Zahedi *et al.* (2014a), who found only 4 per cent of 2011 papers had Tweets (a doubling from 2010 and a quadrupling from 2009). However, the trend identified by both Haustein *et al.* (2014c) and Zahedi *et al.* (2014a) suggests that by 2013 similar samples would have even higher coverage in 2013, in both cases eclipsing the 6 per cent found for the 2013 articles from scielo.br.

For Mendeley, the differences are especially pronounced for the most recent year. The lowest reported coverage in Mendeley (for articles in the most recent year) was 57 per cent (Zahedi *et al.*, 2014a) followed closely by another at 61 per cent (Hammarfelt, 2014). The first of these studies is derived from a random sample chosen from the WoS (from 2011) and the second from a small sample of articles from a Swedish University. For studies that include Mendeley, 60 per cent is in itself a low figure given that most studies report values of over 80 per cent (and some as high as 95 per cent) (Li and Thelwall, 2012; Priem *et al.*, 2012; Bar-Ilan, 2014; Zahedi *et al.*, 2014b). These levels of coverage dwarf the 19 per cent found for scielo.br in 2013, although they are closer to the coverage levels when looking at articles from more than two years prior. The reported values suggest that SciELO Mendeley coverage may approach that found elsewhere, although there appears to be a longer lag in articles being saved into the reference manager (of approximately two years).

The gap between the Facebook coverage of scielo.br and that of previous studies is not as clear. While Priem *et al.* (2012) found a quarter of 2010 PLOS One articles to be shared on Facebook, as few as 2.8 per cent of the 2013 scielo.br articles were mentioned on the platform. This appears to be a significant difference, but a direct comparison is hampered by Altmetric.com's method of querying Facebook, which captures only public posts whereas the method used by Priem *et al.* (2012) captures all posts. As a result, Altmetric's method misses a potentially large segment of all links shared on Facebook. The 2.8 per cent of SciELO articles found on Facebook is consistent with the coverage levels found by Hammarfelt (2014) using Altmetric's data for the small Swedish University sample.

Overall, the difference in the Latin American collections, even in the Brazilian case, indicate that altmetrics show different patterns of coverage than in previously studied contexts (generally lower), especially for the two most prevalent sources (Mendeley and Twitter). Even when compared to the overall coverage of altmetrics measured in previous studies, coverage that includes older articles (which have lower metrics), the contemporary Latin American coverage remains lower. There are at least three plausible explanations for this lower coverage: the articles published in SciELO simply have lower usage overall, and this is reflected in lower social media usage; the levels of social media usage, at least among academics, is lower in Latin America than in the previously studied contexts; or there is a different culture surrounding the sharing of research on social media among the readers of Latin American research. Each of these possible explanations warrants further investigation.

The situation is changing rapidly and will need continuous monitoring. The sharing of research articles on social media appears to be highly variable, making the analysis of trends over time necessary for making sense of what altmetrics have to offer. Furthermore, the evidence presented here suggests that studies of altmetrics cannot be done without the contextualization of the research and social media landscape in which they are embedded. The explanations offered above point to the need of looking at the levels of use of the research, the adoption of social media, and the cultural practices that make up academic work – all of which vary by geography, but also any number of other dimensions, such as discipline and language.

Looking beyond the well-established, high-impact, natural and life science journals is an acknowledgement that “user-generated content [such as what is used for altmetrics] is far from being a simple mirror of either population density or human activity” (Graham *et al.*, 2011, p. 26). Examining the prevalence of altmetrics in different contexts will lead to a better understanding of the significance of the metrics overall, and, at the same time, it will serve to ensure that the eventual use of altmetrics takes into consideration these differences.

Notes

1. There is a parallel here to the underrepresentation of some parts of the world in commercial citation databases that many, including myself, have previously criticized (Alperin, 2014a, b; Cetto and Alonso-Gamboa, 1998; Graham *et al.*, 2011). However, in the case of citation databases, there is at least a known and recognized selection bias. In contrast, the Internet is often seen as a democratizing force, where everyone participates equally. The work of Oxford Internet Institute's Information Geographies project shows that, in practice, this is not the case (Graham *et al.*, 2011, 2014).
2. The terms “Global North/South” are preferred here because they do not have the judgemental connotation of other related terms such as “Developed/Developing World” or

“First/Third World”. The Global North describes the group of countries that are in a privileged economic and political position, such as Canada, the USA, Western Europe, and parts of East Asia.

3. Unsurprisingly, younger scholars seem to use social media more often (Tenopir *et al.*, 2013). However, contrary to expectations, the respondents that report most social media use were also the ones who reported reading the most scholarly papers (Tenopir *et al.*, 2013).
4. The following two recent papers also summarize many of the studies mentioned here and provide some details on sample sizes and correlation coefficients which are omitted here for brevity (Haustein *et al.*, 2014c, Torres-Salinas *et al.*, 2013).
5. These are but a few of the altmetric studies, listed here as examples and not intended to be a comprehensive list. A more complete list of works can be found by looking at the Altmetrics Bibliography (Bailey, 2013) and in the Altmetrics Mendeley group (Taraborelli, n.d.). The altmetrics workshop series (Altmetric.org, 2014) also provides a good overview of the types of studies, and the sources and locations used by the altmetrics community.
6. www.scielo.org
7. SciELO also has national collections in Spain, Portugal, and South Africa, which are not studied here.
8. Details on the APIs can be found at <http://docs.scielo.org/>
9. A detailed description of the extent to which this distribution is representative of all journals in Latin America can be found in Alperin (2014a, b).
10. A detailed description of the sources Altmetric.com tracks can be found on their web site: www.altmetric.com/blog/all-about-altmetrics-sources/
11. All analyses that compare Altmetric.com across national collections use only mentions after November 2013. The analyses specific to SciELO Brazil uses the expanded set starting in January 2013.
12. https://github.com/jalperin/dissertation/blob/master/fetch_mendeley.py
13. https://github.com/jalperin/dissertation/blob/master/fetch_wikipedia.py
14. Sources collected by Altmetric.com which are not listed here, such as Pinterest and LinkedIn, had no mentions.
15. Mendeley and Wikipedia, unlike the Altmetric.com sources, contain all mentions made prior to the collection date. Altmetric.com sources only contain mentions made during the collection window.

References

- Adie, E. and Roe, W. (2013), “Altmetric: enriching scholarly content with article-level discussion and metrics”, *Learned Publishing*, Vol. 26 No. 1, pp. 11-17.
- Alperin, J.P. (2014a), “Assessing the growth and use of open access resources from developing regions: the case of Latin America”, in Alperin, J.P., Babini, D. and Fischman, G. (Eds), *Open Access Indicators and Scholarly Communications in Latin America*, CLACSO, Buenos Aires, pp. 15-78, available at: www.clacso.org.ar/libreria-latinoamericana/libro_detalle.php?orden=&id_libro=906&pageNum_rs_libros=0&totalRows_rs_libros=874 (accessed 9 March 2015).
- Alperin, J.P. (2014b), “South America: citation databases omit local journals”, *Nature*, Vol. 511 No. 7508, p. 155.

- Alperin, J.P., Fischman, G.E. and Willinsky, J. (2011), "Scholarly communication strategies in Latin America's research-intensive universities", *Revista Educación Superior y Sociedad*, Vol. 2 No. 16, pp. 1-9, available at: <http://ess.iesalc.unesco.org.ve/index.php/ess/article/viewArticle/409> (accessed 9 March 2015).
- Altmetrics.org (2014), "Altmetrics14: Expanding Impacts and Metrics", available at: <http://altmetrics.org/altmetrics14/> (accessed 12 December 2014).
- Bailey, C.W. Jr (2013), "Altmetrics bibliography", available at: <http://digital-scholarship.org/alt/altmetrics.htm> (accessed 7 September 2014).
- Bar-Ilan, J. (2012), "JASIST 2001-2010", *Bulletin of the American Society for Information Science and Technology*, Vol. 38 No. 6, pp. 24-28.
- Bar-Ilan, J. (2014), "JASIST@Mendeley revisited", *Altmetrics14: Expanding Impacts and Metrics, Workshop at Web Science Conference 2014, June 23, Bloomington, IN*, available at: http://figshare.com/articles/JASIST_Mendeley_Revisited/1031681 (accessed 9 December 2014).
- Bar-Ilan, J., Haustein, S., Peters, I., Priem, J., Shema, H. and Terliesner, J. (2012), "Beyond citations: scholars' visibility on the social web", *Proceedings of the 17th International Conference on Science and Technology Indicators, Repro-UQAM, Montreal, Canada*, pp. 98-109, available at: http://2012.sticonference.org/Proceedings/vol1/Bar-Ilan_Beyond_98.pdf (accessed 24 January 2014).
- Cetto, A.M. and Alonso-Gamboa, J.O. (1998), "Scientific periodicals in Latin America and the Caribbean: a global perspective", *Interciencia*, Vol. 23 No. 525, pp. 84-93.
- Costas, R., Zahedi, Z. and Wouters, P. (in press), "Do 'altmetrics' correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective", *Journal of the Association for Information Science and Technology*. doi:10.1002/asi.23309.
- Eysenbach, G. (2011), "Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact", *Journal of Medical Internet Research*, Vol. 13 No. 4, p. e123, doi: 10.2196/jmir.2012.
- Fausto, S., Machado, F.A., Bento, L.F.J., Iamarino, A., Nahas, T.R. and Munger, D.S. (2012), "Research blogging: indexing and registering the change in science 2.0.", *PLoS ONE*, Vol. 7 No. 12, p. e50109.
- Galligan, F. and Dias-Correia, S. (2013), "Altmetrics: rethinking the way we measure", *Serials Review*, Vol. 39 No. 1, pp. 56-61.
- Gómez, N.D., Bongiovani, P.C., Moscoloni, N., Rodríguez Jiménez, T.M., de Souza Costa, S.M., Lima Leite, F.C. and De Volder, C. (2013), "Conocimientos y opiniones de los evaluadores investigadores respecto de las publicaciones de Acceso Abierto en Argentina, Brasil y México", available at: <http://es.slideshare.net/nancydianagomez/aa-evaluacion-siaa2013prico4env> (accessed 29 January 2014).
- Graham, M., Hale, S. and Stephens, M. (2011), "Geographies of the world's knowledge", available at: www.oii.ox.ac.uk/publications/convoco_geographies_en.pdf (accessed 16 December 2014).
- Graham, M., Hogan, B., Strauman, R.K. and Medhat, A. (2014), "Uneven geographies of user-generated information: patterns of increasing informational poverty", *Annals of the Association of American Geographers*, Vol. 104 No. 4, pp. 746-764.
- Guédon, J.C. (2008), "Mixing and matching the green and gold roads to open access – take 2", *Serials Review*, Vol. 34 No. 1, pp. 41-51.
- Hammarfelt, B. (2014), "Using altmetrics for assessing research impact in the humanities", *Scientometrics*, Vol. 101 No. 2, pp. 1419-1430.

- Haustein, S., Bowman, T.D., Macaluso, B., Sugimoto, C.R. and Larivière, V. (2014a), "Measuring Twitter activity of arXiv e-prints and published papers", *Altmetrics14: Expanding Impacts and Metrics, Workshop at Web Science Conference 2014, June 23, Bloomington, IN*, available at: <http://dx.doi.org/10.6084/m9.figshare.1041514> (accessed 9 December 2014).
- Haustein, S., Peters, I., Sugimoto, C.R., Thelwall, M. and Larivière, V. (2014c), "Tweeting biomedicine: an analysis of tweets and citations in the biomedical literature", *Journal of the Association for Information Science and Technology*, Vol. 65 No. 4, pp. 656-669.
- Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H. and Terliesner, J. (2014b), "Coverage and adoption of altmetrics sources in the bibliometric community", *Scientometrics*, Vol. 101 No. 2, pp. 1145-1163.
- Li, X. and Thelwall, M. (2012), "F1000, Mendeley and traditional bibliometric indicators", *Proceedings of the 17th International Conference on Science and Technology Indicators, Repro-UQAM, Montreal, Canada*, pp. 541-551, available at: http://2012.sticonference.org/Proceedings/vol2/Li_F1000_541.pdf (accessed 26 January 2014).
- Li, X., Thelwall, M. and Giustini, D. (2012), "Validating online reference managers for scholarly impact measurement", *Scientometrics*, Vol. 91, No. 2, pp. 461-471.
- Liu, J. and Adie, E. (2013), "Altmetric: getting started with article-level metrics", figshare, 30 May, doi:10.6084/m9.figshare.709018.
- NISO (2014), "Alternative metrics initiative Phase 1", White Paper, NISO", available at: www.niso.org/apps/group_public/download.php/13809/Altmetrics_project_phase1_white_paper.pdf (accessed 9 March 2015).
- Packer, A.L. and Meneghini, R. (2007), "Learning to communicate science in developing countries", *Interciencia*, Vol. 32 No. 9, pp. 643-647.
- Piwowar, H. (2013), "Introduction altmetrics: what, why and where?", *Bulletin of the American Society for Information Science and Technology*, Vol. 39 No. 4, pp. 8-9.
- Priem, J., Piwowar, H. and Hemminger, B.M. (2012), "Altmetrics in the wild: using social media to explore scholarly impact", available at: <http://arxiv.org/abs/1203.4745> (accessed 22 March 2012).
- Procter, R., Williams, R. and Stewart, J. (2010), "If you build it, will they come? How researchers perceive and use web 2.0", available at: http://wrap.warwick.ac.uk/56246/1/WRAP_Procter_If%20you%20build%20it%20will%20they%20come.pdf (accessed 17 December 2014).
- Shema, H., Bar-Ilan, J. and Thelwall, M. (2012), "Research blogs and the discussion of scholarly information", *PLOS ONE*, Vol. 7 No. 5, p. e35869.
- Shuai, X., Pepe, A. and Bollen, J. (2012), "How the scientific community reacts to newly submitted preprints: article downloads, Twitter mentions, and citations", in Ouzounis, C.A. (Ed.), *PLoS One*, Public Library of Science, Vol. 7 No. 11, p. e47523.
- Taraborelli, D. (n.d.), "Altmetrics Mendeley Group", available at: www.mendeley.com/groups/586171/altmetrics/ (accessed 12 December 2014).
- Tenopir, C., Volentine, R. and King, D.W. (2013), "Social media and scholarly reading", *Online Information Review*, Vol. 37 No. 2, pp. 193-216.
- Thelwall, M., Haustein, S., Larivière, V. and Sugimoto, C. (2013), "Do altmetrics work? Twitter and ten other social web services", *PLOS ONE*, Vol. 8 No. 5, p. e64841.
- Torres-Salinas, D., Cabezas-Clavijo, Á. and Jiménez-Contreras, E. (2013), "Altmetrics: new indicators for scientific communication in web 2.0", *Comunicar*, Vol. 21 No. 41, pp. 53-60.
- Wouters, P. and Costas, R. (2012), "Users, narcissism and control: tracking the impact of scholarly publications in the 21st century", *Proceedings of the 17th International Conference on Science and Technology Indicators, Repro-UQAM, Montreal, Canada*, pp. 847-857, available

at: http://2012.sticonference.org/Proceedings/vol2/Wouters_Users_847.pdf (accessed 25 January 2014).

Zahedi, Z., Costas, R. and Wouters, P. (2014a), "How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications", *Scientometrics*, Vol. 101 No. 2, pp. 1491-1513.

Zahedi, Z., Fenner, M. and Costas, R. (2014b), "How consistent are altmetrics providers? Study of 1000 PLOS ONE publications using the PLOS ALM, Mendeley and Altmetric.com APIs", *Altmetrics14: Expanding Impacts and Metrics, Workshop at Web Science Conference 2014, June 23, Bloomington, IN*, available at: <http://dx.doi.org/10.6084/m9.figshare.1041821> (accessed 10 December 2014).

Corresponding author

Juan Pablo Alperin can be contacted at: juan@alperin.ca

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com